

# Polymorphic Internal Transcribed Spacer Region 1 DNA Sequences Identify Medically Important Yeasts

YI-CHING CHEN,<sup>1</sup> JESSICA D. EISNER,<sup>1</sup> MIREILLE M. KATTAR,<sup>1</sup> SARA L. RASSOULIAN-BARRETT,<sup>1</sup>  
KAREN LAFE,<sup>1</sup> UYEN BUI,<sup>1</sup> AJIT P. LIMAYE,<sup>1,2</sup> AND BRAD T. COOKSON<sup>1,3\*</sup>

*Departments of Laboratory Medicine,<sup>1</sup> Infectious Diseases,<sup>2</sup> and Microbiology,<sup>3</sup>  
University of Washington, Seattle, Washington 98195*

Received 11 June 2001/Returned for modification 15 August 2001/Accepted 28 August 2001

**Species-specific polymorphisms in the noncoding internal transcribed spacer 2 (ITS2) region of the rRNA operon provide accurate identification of clinically significant yeasts. In this study, we tested the hypothesis that ITS1 noncoding regions contain diagnostically useful alleles. The length of ITS1 region PCR products amplified from 40 species (106 clinical strains, 5 reference strains, and 30 type strains) was rapidly determined with single-base precision by automated capillary electrophoresis. Polymorphisms in the PCR product length permitted 19 species to be distinguished by ITS1 alone, compared with 16 species distinguished by using only ITS2. However, combination of both ITS alleles permitted identification of 30 species (98% of clinical isolates). The remaining 10 species with PCR products of similar sizes contained unique ITS alleles distinguishable by restriction enzyme analysis. DNA sequence analysis of amplified ITS1 region DNA from 79 isolates revealed species-specific ITS1 alleles for each of the 40 pathogenic species examined. This provided identification of unusual clinical isolates, and 53 diagnostic ITS1 sequences were deposited in GenBank. Phylogenetic analyses based on ITS sequences showed a similar overall topology to 26S rRNA gene-based trees. However, different species with identical 26S sequences contained distinct ITS alleles that provided species identification with strong statistical support. Together, these data indicate that the analysis of ITS polymorphisms can reliably identify 40 species of clinically significant yeasts and that the capacity for identifying potentially new pathogenic species by using this database holds significant promise.**

The complexity of opportunistic fungal infections is increasing as more patients are adversely affected by an expanding diversity of yeasts. Invasive procedures or immunosuppression increases the risk of fungemia (24, 34), and drug-resistant strains have emerged (9, 21). Although *Candida albicans* is still the predominate agent of nosocomial infection, serious infections caused by other yeasts have increased in frequency (8, 11, 29, 34). For example, species of *Cryptococcus* other than *Cryptococcus neoformans*, previously considered to be nonpathogenic saprophytes, have been reported to cause cryptococcosis (28). Adequate treatment of these infections depends on early detection and accurate identification of the pathogens. However, conventional identification by evaluation of morphological and physiological characteristics can be laborious, sometimes leads to incorrect classification and identification (2, 3, 13), and can be impeded by database limitations (4, 6).

To develop a rapid molecular method for identifying yeasts, we recently analyzed the length and sequence polymorphisms in the DNA of noncoding internal transcribed spacer region 2 (ITS2) of the rRNA operon (2). ITS2 region polymorphisms permitted accurate identification of over 400 clinical strains representing 34 species of pathogenic yeasts. ITS2 region PCR product length determined by automated capillary electrophoresis permitted single-base-pair precision, with between-run standard deviations (SD) of  $\leq 0.5$  base, and 92% of the clinical strains examined were rapidly and correctly identified

by using only the unique length of their PCR products. The remaining 8% could be identified by either restriction endonuclease or DNA sequence analysis of their ITS2 region PCR products.

In this study, we investigated the length and sequence polymorphisms in the ITS1 region and determined their diagnostic and phylogenetic utility for identification of medically important yeasts. These DNA sequences contained unique alleles for all 40 pathogenic species examined, and ITS1 alleles displayed greater interspecies variability than ITS2 region sequences. Thirty species, comprising 98% of the clinical isolates, were easily distinguished by simply determining the length of the ITS1 and ITS2 region PCR products. The remaining 10 species could be differentiated by either restriction endonuclease or DNA sequence analysis of ITS PCR products. The sequence diversity among ITS1 alleles was useful for distinguishing closely related species, particularly those that contained identical DNA sequences in the D1/D2 variable domain of the 26S rRNA gene. Thus, our ITS sequence data provide a reliable means with which to rapidly identify known yeast pathogens, will facilitate the discovery of potentially novel pathogens, and establish a foundation for further expansion of an ITS sequence database of medically important fungi.

## MATERIALS AND METHODS

**Yeast isolates.** We examined 106 clinical strains and five reference strains (Tables 1 and 2) previously characterized genetically, morphologically, and biochemically in detail (2). These strains were from our clinical laboratory at the University of Washington Medical Center. Twenty-three type strains from the American Type Culture Collection (ATCC) and 7 type strains from the Centraalbureau voor Schimmelfcultures (CBS) were also included in this study (Table 2). These 141 strains represented 40 different species of pathogenic yeasts.

\* Corresponding author. Mailing address: Departments of Laboratory Medicine and Microbiology, University of Washington, 1959 NE Pacific St., NW 120, Box 357110, Seattle, WA 98195. Phone: (206) 598-6131. Fax: (206) 598-6189. E-mail: cookson@u.washington.edu.

TABLE 1. Lengths of ITS1 and ITS2 region PCR products from clinical isolates and type strains

Species <sup>a</sup> (no. of isolates)	ITS1 PCR product length <sup>b</sup> (mean no. of bp)	SD <sup>c</sup> (bp)	Range of product lengths (bp)		ITS2 PCR product length <sup>d</sup> (mean no. of bp)
			Minimum	Maximum	
<i>Yarrowia lipolytica</i> (2)	134.43	0.29	134.10	134.60	236.66
<i>Candida rugosa</i> <sup>T</sup> (1)	142.91		NA <sup>e</sup>	NA	270.99
<i>Candida lusitanae</i> (4)	143.87	0.35	143.54	144.44	251.14
<i>Candida intermedia</i> (1)	148.48		NA	NA	255.10
<i>Pichia ohmeri</i> <sup>T</sup> (1)	154.01		NA	NA	268.84
<i>Candida lambica</i> (1)	156.39		NA	NA	302.18
<i>Candida pararugosa</i> (1) <sup>f</sup>	161.34		NA	NA	269.57
<i>Candida krusei</i> (5)	181.44	0.64	180.39	182.27	344.61
<i>Cryptococcus laurentii</i> <sup>T</sup> (1)	190.84		NA	NA	362.83
<i>Trichosporon jirovecii</i> <sup>T</sup> (1)	192.33		NA	NA	349.89
<i>Trichosporon cutaneum</i> <sup>T</sup> (1)	193.72		NA	NA	350.00
<i>Trichosporon mucoides</i> (1)	194.25		NA	NA	349.30
<i>Cryptococcus neoformans</i> (5)	197.93	0.26	197.59	198.47	370.76
<i>Cryptococcus humicola</i> <sup>T</sup> (1)	197.94		NA	NA	353.17
<i>Trichosporon inkin</i> (1)	197.97		NA	NA	353.27
<i>Trichosporon asahii</i> (1)	200.00		NA	NA	355.87
<i>Cryptococcus uniguttulatus</i> (1)	208.00		NA	NA	428.73
<i>Candida utilis</i> <sup>T</sup> (1)	213.92		NA	NA	362.34
<i>Candida dubliniensis</i> (7) <sup>g</sup>	214.06	0.46	213.40	214.81	342.33
<i>Candida albicans</i> (8)	214.35	0.41	214.00	214.99	339.34
<i>Candida tropicalis</i> (11)	214.78	0.45	213.95	215.73	327.42
<i>Candida freyschussii</i> <sup>T</sup> (1)	223.46		NA	NA	347.27
<i>Candida parapsilosis</i> (4)	225.43	0.11	225.30	225.58	309.93
<i>Sporobolomyces salmonicolor</i> <sup>T</sup> (1)	227.45		NA	NA	400.19
<i>Rhodotorula rubra</i> (7)	228.36	0.33	228.41	229.41	400.46
<i>Rhodotorula glutinis</i> <sup>T</sup> (1)	229.92		NA	NA	390.93
<i>Cryptococcus albidus</i> <sup>T</sup> (1)	231.03		NA	NA	403.67
<i>Cryptococcus liquefaciens</i> (1) <sup>h</sup>	231.39		NA	NA	403.45
<i>Cryptococcus diffluens</i> (1) <sup>h</sup>	231.41		NA	NA	403.75
<i>Candida guilliermondii</i> (5)	243.75	0.21	243.51	244.11	374.61
<i>Hansenula anomala</i> (3)	257.35	0.12	257.21	257.53	372.58
<i>Pichia veronae</i> <sup>T</sup> (1)	261.44		NA	NA	365.95
<i>Pichia fabianii</i> (1) <sup>f</sup>	265.45		NA	NA	364.78
<i>Candida zeylanoides</i> (2)	267.72	0.29	267.27	268.03	371.65
<i>Candida famata</i> (2)	273.52	0.21	273.32	273.73	376.21
<i>Endomyces fibuliger</i> (2)	291.13	1.14	289.56	292.25	374.98
<i>Pichia farinosa</i> (2)	303.19	0.00	303.18	303.19	379.18
<i>Candida kefyr</i> (3)	305.45	0.15	305.31	305.60	427.88
<i>Saccharomyces cerevisiae</i> (6)	436.52	2.09	434.75	440.03	416.35
<i>Candida glabrata</i> (17)	475.24	0.29	474.78	475.68	413.51

<sup>a</sup> Identification of species by biochemical, morphological, and genotypic assessment. The number of isolates analyzed is indicated in parentheses. Type strains are designated by a superscript T.

<sup>b</sup> PCR product sizes determined by capillary electrophoresis as described in Materials and Methods.

<sup>c</sup> SD are calculated for species with two or more strains.

<sup>d</sup> ITS2 PCR product lengths are from our previous study (2).

<sup>e</sup> NA, nonapplicable.

<sup>f</sup> *Candida pararugosa* and *Pichia fabianii* were designated as isolate UWFP-348 and UWFP-345, respectively, in a previous study (2).

<sup>g</sup> Three strains obtained from the University of Texas Health Science Center at San Antonio (2).

<sup>h</sup> *C. liquefaciens* UWFP-357 and *C. diffluens* UWFP-359 were both misidentified as *Cryptococcus albidus* by the VITEK and API systems (2).

Nitrate assimilation, which distinguishes *Pichia fabianii* from *Pichia veronae*, was performed on a slant of nitrate agar (1.4 g of potassium nitrate, 1.6 g of yeast carbon base from Difco, 0.12 g of bromthymol blue, 16 g of Noble agar, 1 liter of distilled water [pH 5.9 to 6.0]) incubated at 30°C for about 1 week until the agar turned from yellow or yellow-green (negative) to blue (positive). *Cryptococcus albidus* and *Candida albicans* were used as positive and negative controls, respectively (17).

A hierarchical approach to building a cross-validated data set was employed when selecting isolates for molecular analyses. This was based on our observations that (i) congeneric strains identified according to their genetic, biochemical, and morphological characters generally show >99% sequence similarity at ITS2 (2) and ITS1 (Table 2); (ii) capillary electrophoresis permits single-base-pair precision in determining the length of PCR products, with SD of  $\leq 0.5$  base between runs; and (iii) the apparent PCR product length strongly correlates ( $R^2$

TABLE 2. Length and sequence polymorphisms of ITS1 region DNAs from clinical and type strains

Species	Strain(s) <sup>a</sup>	ITS1 PCR product length <sup>b</sup> (bp)	Sequenced PCR product <sup>c</sup> (bp)	% Identity with type strains	GenBank accession no.	Source or reference
<i>Candida albicans</i>	ATCC 11006	215.09	219		AF336831 <sup>d</sup>	TS <sup>k</sup>
	UWFP-60	214.17	218	99.5	AF336832 <sup>d</sup>	TS
	ATCC 28516	NT <sup>i</sup>	218	99.5	AF217609	
<i>Candida dubliniensis</i>	UWFP-92	214.06	218		AF336833	TS
	CD36	NT	218	100 <sup>e</sup>	AJ249485	
<i>Candida famata</i>	CBS1795 <sup>T</sup>	273.63	278		AF336834	TS
	UWFP-352	273.48	278	100	AF336834	TS
	CBS767	NT	278	100	AF210327	
<i>Candida freyschussii</i>	ATCC 18737 <sup>T</sup>	223.46	231		AF336835	TS
<i>Candida glabrata</i>	ATCC 2001 <sup>T</sup>	475.50	482		AF336836	TS
	UWFP-115, -118	475.24	482	99.8	AF336837	TS
	UWFP-192	475.31	482	99.8	AF336838	TS
	IFO 0622	NT	452	100	AB032177	
<i>Candida guilliermondii</i>	ATCC 6260 <sup>T</sup>	243.08	248		AF336839	TS
	UWFP-195	243.73	248	100	AF336839	TS
	Taxon:4929	NT	249	99.6	L47110	33
<i>Candida intermedia</i>	ATCC 14439 <sup>T</sup>	147.42	151		AF336840	TS
	UWFP-347	148.48	151	100	AF336840	TS
<i>Candida kefyr</i>	UWFP-208	305.60	309		AF336841	TS
	JCM1630	NT	279	99.7 <sup>f</sup>	AB011519	22
<i>Candida krusei</i>	ATCC 6258 <sup>T</sup>	180.41	182		AF411417	TS
<i>Candida krusei</i> sequevar 1	UWFP-210, -211	181.87	183	98.4 <sup>e</sup>	AF336842	TS
	ATCC 6258 <sup>T</sup>	NT	182	100	AF246989	
<i>Candida lambica</i>	ATCC 24750 <sup>T</sup>	156.42	158		AF336843	TS
	UWFP-346	156.39	158	100	AF336843	TS
<i>Candida lusitanae</i>	ATCC 34449 <sup>T</sup>	144.34	148		AF336844	TS
	UWFP-233, -234	144.07	148	100	AF336844	TS
	ATCC 34449 <sup>T</sup>	NT	148	98.0	AF009215	
<i>Candida parapsilosis</i>	UWFP-254, -285	225.44	229		AF336846	TS
	ATCC 22019	NT	229	100 <sup>e</sup>	U10987	16
<i>Candida pararugosa</i>	ATCC 38774 <sup>T</sup>	161.37	164		AF411418	TS
	UWFP-348	161.34	164	99.4	AF335925 <sup>d</sup>	TS
<i>Candida pelliculosa</i> ( <i>Hansenula anomala</i> )	CBS605 <sup>T</sup>	257.34	262		AF335926	TS
	UWFP-396	257.36	262	100	AF335926	TS
	FY-102	NT	262	100	AF270936	19
<i>Candida rugosa</i>	ATCC 10571 <sup>T</sup>	142.91	145		AF335927	TS
<i>Candida tropicalis</i>	ATCC 750 <sup>T</sup>	214.83	218		AF335928	TS
	UWFP-313, -321, -324	214.94	218	100	AF335928	TS
	Taxon:5482	NT	218	99.5	L47112	33
<i>Candida utilis</i>	ATCC 22023 <sup>T</sup>	213.92	220		AF335929	TS
<i>Candida zeylanoides</i>	ATCC 7351 <sup>T</sup>	267.81	272		AF335930	TS
	UWFP-349	267.64	272	100	AF335930	TS
<i>Cryptococcus albidus</i>	ATCC 10666 <sup>T</sup>	231.03	234		AF335931 <sup>g</sup>	TS
	CBS142 <sup>T</sup>	NT	234	100	AF145321	26
<i>Cryptococcus diffluens</i>	UWFP-359	231.41	234		AF335932	TS
	CBS160 <sup>T</sup>	NT	234	100 <sup>e</sup>	AF145330	26
<i>Cryptococcus humicolus</i>	ATCC 14438 <sup>T</sup>	197.94	200		AF335933	TS
	UWFP-362	197.91	200	99.5	AF335934	TS
	JCM1457	NT	170	100 <sup>h</sup>	AB035572	29
<i>Cryptococcus laurentii</i>	ATCC 18803 <sup>T</sup>	190.84	193		AF335935	TS
	JCM9066	NT	163	99.4 <sup>h</sup>	AB035043	28
<i>Cryptococcus liquefaciens</i>	UWFP-357	231.39	234		AF336845 <sup>g</sup>	TS
<i>Cryptococcus neoformans</i>	ATCC 32045 <sup>T</sup>	198.20	201		AF335936	TS
	UWFP-360	197.74	201	99.0	AF335937	TS
	Taxon:37769	NT	201	99.0	AF196312	

Continued on following page

TABLE 2—Continued

Species	Strain(s) <sup>a</sup>	ITS1 PCR product length <sup>b</sup> (bp)	Sequenced PCR product <sup>c</sup> (bp)	% Identity with type strains	GenBank accession no.	Source or reference
<i>Cryptococcus uniguttulatus</i>	CBS1730 <sup>T</sup>	207.71	212		AF335938	TS
	UWFP-364	208.00	212	100	AF335938	TS
	JCM3685	NT	212	100	AB032692	31
<i>Endomyces fibuliger</i>	CBS329.83 <sup>T</sup>	290.92	296		AF335939	TS
	UWFP-397	290.62	295	99.7	AF335940	TS
	UWFP-398	291.64	297	99.7	AF335941	TS
	8014	NT	296	100	U10409	
<i>Pichia fabianii</i>	CBS5640 <sup>T</sup>	265.62	270		AF335942 <sup>d</sup>	TS
	UWFP-345	265.45	270	100	AF335942	TS
<i>Pichia farinosa</i>	CBS185 <sup>T</sup>	304.46	307		AF335944	TS
<i>Pichia farinosa</i> sequevar 1	UWFP-389, -390	303.19	307	98.4	AF335945	TS
<i>Pichia ohmeri</i>	ATCC 46053 <sup>T</sup>	154.01	156		AF335946	TS
<i>Pichia ohmeri</i> sequevar 1	UWFP-388	155.48	158	88.9	AF335947 <sup>e</sup>	TS
<i>Pichia veronae</i>	CBS6591 <sup>T</sup>	261.44	265		AF335943 <sup>d</sup>	TS
<i>Rhodotorula glutinis</i>	ATCC 32765 <sup>T</sup>	229.92	232		AF335948	TS
	JCM8208	NT	232	98.3	AB026018	
<i>Rhodotorula rubra</i>	ATCC 32763 <sup>T</sup>	229.53	232		AF335949	TS
	UWFP-370, -371, -374	228.80	232	100 <sup>f</sup>	AF335950 <sup>d</sup>	TS
	UWFP-373, -380	228.61	232	99.1 <sup>i</sup>	AF335951 <sup>d</sup>	TS
	SY-93	NT	232	100	AB026003	
<i>Saccharomyces cerevisiae</i>	UWFP-382, -384	435.09	441	99.8 <sup>j</sup>	AF335952	TS
	UWFP-383	439.90	446	98.7 <sup>j</sup>	AF335953	TS
	UWFP-387	437.11	443	99.3 <sup>j</sup>	AF335954	TS
	CBS4903	NT	424 <sup>g</sup>		Z95940	
<i>Sporobolomyces salmonicolor</i>	ATCC 36400 <sup>T</sup>	227.45	231		AF335955	TS
	JCM1841	NT	231	98.7	AB030341	30
<i>Trichosporon asahii</i>	UWFP-391	200.00	203		AF335956	TS
	CBS2530	NT	173	100 <sup>f</sup>	AB018014	27
<i>Trichosporon cutaneum</i>	ATCC 28592 <sup>T</sup>	193.72	196		AF335957	TS
	CBS2466 <sup>T</sup>	NT	166	100 <sup>h</sup>	AB018020	27
<i>Trichosporon inkin</i>	ATCC 18020 <sup>T</sup>	198.47	201		AF335958	TS
	UWFP-399	197.97	201	100	AF335958	TS
	CBS5585 <sup>T</sup>	NT	170	99.4 <sup>h</sup>	AB018024	27
<i>Trichosporon jirovecii</i>	ATCC 34499 <sup>T</sup>	192.33	195		AF335959	TS
	CBS6950	NT	165	100 <sup>h</sup>	AB018026	27
<i>Trichosporon mucoides</i>	UWFP-363, -366, -367	194.37	197		AF335960 <sup>d</sup>	TS
	M9478	NT	167	100 <sup>f</sup>	AB018031	27
<i>Yarrowia lipolytica</i>	ATCC 18942 <sup>T</sup>	133.97	139		AF335961	TS
<i>Y. lipolytica</i> sequevar 1	UWFP-400, -401	134.48	139	97.8	AF335962	TS

<sup>a</sup> ATCC, American Type Culture Collection; UWFP, University of Washington Fungal Project; CD, Cardiff Dental; CBS, Centraalbureau voor Schimmelcultures; IFO, Institute of Fermentation; JCM, Japan Collection of Microorganisms; SY, deep sea yeast; M, Meiji Pharmaceutical; type strains are labeled with superscript T.

<sup>b</sup> PCR product length determined with capillary electrophoresis as described in Materials and Methods.

<sup>c</sup> Exact number of nucleotides determined by direct sequencing of the PCR products.

<sup>d</sup> In addition to ITS1 data listed, new ITS2 sequence data include *C. albicans* ATCC 11006 (AF335963), *C. albicans* UWFP-60 (AF335964), *C. pararugosa* UWFP-348 (AF335965), *P. fabianii* CBS5640<sup>T</sup> (AF335967), *P. veronae* CBS6591<sup>T</sup> (AF335966), *R. rubra* UWFP-370 (AF335969), *R. rubra* UWFP-373 (AF335968), and *T. mucoides* UWFP-366 (AF335970).

<sup>e</sup> Complete sequence compared to ITS1 region DNA sequences from clinical strains in this study.

<sup>f</sup> Partial sequence compared to ITS1 region DNA sequences from clinical strains in this study.

<sup>g</sup> UWFP-357 and -359 were misidentified as *Cryptococcus albidus* by the VITEK system (2). Their ITS2 sequences were identical and submitted as *C. albidus* under accession no. AF219002.

<sup>h</sup> Partial sequence compared to ITS1 region DNA sequences from type strains in this study.

<sup>i</sup> UWFP-373 and -380 had identical ITS2 sequences, and they were 99.3% similar to the *R. rubra* type strain. UWFP-370 had 99.5% sequence similarity to the *R. rubra* type strain.

<sup>j</sup> ITS1 sequence similarities between UWFP-383 and -387, UWFP-382 and -383, and UWFP-382 and -387 were 98.4, 98.9, and 99.5%, respectively.

<sup>k</sup> TS, this study.

<sup>l</sup> NT, not tested.

TABLE 3. Species with similar ITS region PCR product lengths are distinguishable by species-specific ITS1 and ITS2 DNA sequences or restriction fragments of the ITS1-ITS2-region

Species	Length (bp)		% Sequence similarity		Restriction fragment size(s) <sup>b</sup> (bp)	
	ITS1	ITS2 <sup>a</sup>	ITS1	ITS2 <sup>a</sup>	<i>Bsa</i> HI	<i>Hinc</i> II
<i>T. jirovecii</i> <sup>T</sup> (A)	192.33	349.89	A	A	527	49/478
<i>T. cutaneum</i> <sup>T</sup> (B)	193.72	350.00	98.0, B	97.7, B	529	53/74/102
<i>T. mucoides</i> (C)	194.25	349.30	97.0, 97.5	99.4, 97.2	528	528
<i>C. humicolus</i> <sup>T</sup>	197.94	353.17	84.6	85.2	535	535
<i>T. inkin</i>	197.97	353.27			350/189	539
<i>C. dubliniensis</i>	214.06	342.33 <sup>c</sup>	93.2	92.1	457/84	541
<i>C. albicans</i>	215.09	338.63 <sup>c</sup>			536	536
<i>S. salmonicolor</i> <sup>T</sup>	227.45	400.19	80.6	83.7	612	612
<i>R. rubra</i> <sup>T</sup>	229.53	400.00			41/575	616
<i>C. albidus</i> <sup>T</sup> (A)	231.03	403.67	A	A	621	621
<i>C. liquefaciens</i> (B)	231.39	403.45	96.2, B	98.0, B	621	363/258
<i>C. diffluens</i> (C)	231.41	403.75	96.6, 99.6	98.0, 100	621	363/258

<sup>a</sup> Data from reference 2.<sup>b</sup> Restriction enzyme analyses of PCR products containing ITS1 plus 5.8S rRNA gene plus ITS2 sequences.<sup>c</sup> These isolates can be distinguished by using ITS2 length polymorphisms (Table 1) but are included here because of the difficulty of phenotypic identification (2).

= 0.9992) with the actual number of nucleotides enumerated by direct sequencing (2). Type and reference strains were characterized by directly sequencing ITS1 region DNA. ITS1 PCR product length polymorphisms were determined for randomly selected representatives of each of 40 species (Table 1): for 18 species, multiple isolates were characterized, and for 22 species, only single isolates were characterized for PCR product length. All length polymorphisms (Table 1) were subsequently confirmed by direct sequence analysis of, in most cases, more than one representative isolate of each species (Table 2). Among the 22 singly listed isolates in Table 1, 7 isolates represent type strains added to this study for completeness and for which no clinical isolates were available, 7 isolates showed >99% sequence similarity with their respective type strains, 7 isolates showed >99% sequence similarity with another sequenced clinical isolate, and 1 clinical isolate, *Cryptococcus liquefaciens*, was the only representative of that species characterized by analyzing DNA sequences from ITS1 (Table 2) and ITS2 (Table 3) and the D1/D2 variable domain of the 26S rRNA gene (Table 4).

**PCR and DNA sequencing.** DNA was extracted from yeasts after 48 h of growth on Sabouraud's agar and PCR amplified as previously described (2). ITS1 region DNA was amplified with 900 nM primer ITS1 (5'-TCCGTAGGTGAAC CTGCGG-3'; GIBCO BRL, Grand Island, N.Y.), 300 nM primer ITS2 (5'-GC TGCGTCTTCATCGATGC-3') (18), and the following thermocycler parameters: 95°C for 6 min, followed by 25 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s, followed by one final extension at 72°C for 10 min. The same parameters were used to amplify the ITS1 and ITS2 regions simultaneously as a single product by using 300 nM ITS1 and ITS4 primers (5'-TCCTCCGCTTAT TGATATGC-3') (18). Similarly, 600 nM (each) NL-1 (5'-GCATATCAATAA GCGGAGGAAAAG-3') and NL-4 (5'-GGTCCGTGTTCAAGACGG-3') (14) was used to amplify the D1/D2 variable domain of the 26S ribosomal DNA (rDNA) gene: denaturation at 94°C for 5 min was followed by 30 cycles of amplification at 94°C for 1 min, 60°C for 2 min, and 72°C for 2 min, followed by one final extension at 72°C for 7 min. Determination of the length of ITS1 PCR products by automated capillary electrophoresis and DNA sequence analysis of both the forward and reverse strands of PCR products (listed in Table 2) were performed as previously described (2).

**Sequence similarity and phylogenetic analyses.** DNA sequences were assembled, edited, and subjected to phylogenetic analyses as described previously (2). Sequence similarities were expressed as the percentage of nucleotide differences determined by pairwise sequence comparisons. *Saccharomyces cerevisiae* was used as the outgroup instead of *Pneumocystis carinii* (2): the ITS1 sequence of *S. cerevisiae* is distant from all the other sequences (except *Candida glabrata*) and thus forms a natural outgroup. The branching order of the neighbor-joining (23) dendrograms (Fig. 1) was evaluated with 1,000 bootstrap analyses by using the SEQBOOT program in the PHYLIP software package (version 3.573) (J. Felsenstein, Department of Genetics, University of Washington, Seattle; <http://evolution.genetics.washington.edu/phytip.html>).

## RESULTS

To determine if DNA polymorphisms in the ITS1 region contained diagnostically useful information, we analyzed single PCR products that had been amplified from 141 strains of yeasts, including 30 type and 5 reference strains. Each product contained the 3' end of the 18S rDNA gene, the entire ITS1 region, and the 5' end of the 5.8S rDNA gene. The PCR product length was analyzed for each strain by capillary electrophoresis under denaturing conditions (Table 1), a method that rapidly provides single-base-pair precision, low run-to-run variation, and an excellent correlation ( $R^2 = 0.9992$ ) with the actual number of nucleotides enumerated by direct sequencing (2). To confirm the specificity of length polymorphisms and identify species-specific ITS1 alleles, the forward and reverse strands of ITS1 region PCR products from 79 strains were directly sequenced (Table 2).

**ITS1 sequence polymorphisms.** ITS1 PCR product lengths determined by capillary electrophoresis ranged from 134.43 bp (*Yarrowia lipolytica*) to 475.24 bp (*C. glabrata*) with intraspecies strain variation generally less than 0.5 bp, except for *Candida krusei* (SD = 0.64 bp), *Endomyces fibuliger* (SD = 1.14 bp), and *S. cerevisiae* (SD = 2.09 bp) (Table 1). Similar intraspecies strain variation was observed among ITS2 alleles for *Candida dubliniensis*, *Candida tropicalis*, and *S. cerevisiae* (2), but the distribution of ITS2 region PCR product lengths among all tested species was more limited and ranged from 236.66 to 428.73 bp (Table 1). Of 40 species examined, 19 could be identified based solely on the apparent length of their ITS1 PCR products (species listed individually in Table 1).

Direct DNA sequence analysis of the ITS1 region PCR products (Table 2) confirmed a close correlation of the apparent ITS1 PCR product lengths measured by capillary electrophoresis with the actual product length (2) and revealed species-specific ITS1 DNA sequences for all 40 pathogenic species examined. Intraspecies sequence variability was detected among clinical strains and between clinical and type strains



TABLE 4. Comparing three diagnostic loci for identifying unusual isolates

Group and strain(s) <sup>a</sup>	Identity by biochemical tests <sup>b</sup>	Actual identity by genetic analyses <sup>c</sup>	Sequence similarity among clinical and type strains (%) <sup>d</sup>			Accession no. from this study for 26S
			ITS1	ITS2	26S	
Group 1 UWFP-208	<i>C. kefir</i>	<i>C. kefir</i>	99.7	98.4	100	AF335978
Group 2 ATCC 6258 <sup>T</sup> UWFP-210, -211	<i>C. krusei</i> <i>C. krusei</i>	<i>C. krusei</i> <i>C. krusei</i>	98.4	100	100	AF335979
Group 3 ATCC 10571 <sup>T</sup> UWFP-348 ATCC 38774 <sup>T</sup>	<i>C. rugosa</i> <i>C. rugosa</i> <i>C. pararugosa</i>	<i>C. rugosa</i> (A) <i>C. pararugosa</i> (B) <i>C. pararugosa</i> (C)	A 69.5, B 68.9, 99.4	A 62.8, B 61.1, 100	A 63.7, B 63.7, 100	AF335972 AF421856
Group 4 ATCC 10666 <sup>T</sup> UWFP-359 UWFP-357	<i>C. albidus</i> <i>C. albidus</i> <i>C. albidus</i>	<i>C. albidus</i> (A) <i>C. diffluentis</i> (B) <i>C. liquefaciens</i> (C)	A 96.6, B 96.2, 99.6	A 98.0, B 98.0, 100	A 97.0, B 97.3, 99.5	AF335982 AF335981 AF335980
Group 5 ATCC 32045 <sup>T</sup> UWFP-360	<i>C. neoformans</i> <i>C. neoformans</i>	<i>C. neoformans</i> <i>C. neoformans</i>	99.0	99.0	100 99.5	AF335984 AF335983
Group 6 CBS185 <sup>T</sup> UWFP-389, -390	<i>P. farinosa</i> <i>P. farinosa</i>	<i>P. farinosa</i> <i>P. farinosa</i> sequevar 1	98.4	100	100 99.7	AF335974 AF335973
Group 7 ATCC 46053 <sup>T</sup> UWFP-388	<i>P. ohmeri</i> <i>P. ohmeri</i>	<i>P. ohmeri</i> <i>P. ohmeri</i> sequevar 1	88.9	99.6	100 100	AF335976 AF335975
Group 8 ATCC 32765 <sup>T</sup> ATCC 32763 <sup>T</sup> UWFP-370 UWFP-373, -380	<i>R. glutinis</i> <i>R. rubra</i> <i>R. glutinis</i> <i>R. glutinis</i>	<i>R. glutinis</i> (A) <i>R. rubra</i> (B) <i>R. rubra</i> (C) <i>R. rubra</i> (D)	A 90, B 90, 100, C 90, 99, 99	A 91, B 90, 99, C 91, 99, 100	A 94.3, B 94.3, 100, C 94.3, 99.8, 99.8	AF335985 AF335986 AF335986 AF335987
Group 9 ATCC 18942 <sup>T</sup> UWFP-400, -401	<i>Y. lipolytica</i> <i>Y. lipolytica</i>	<i>Y. lipolytica</i> <i>Y. lipolytica</i> sequevar 1	97.8	100	100 100	AF335977 AF335977
Group 10 UWFP-345 CBS5640 <sup>T</sup> CBS6591 <sup>T</sup>	Unknown <sup>e</sup> <i>P. fabianii</i> <i>P. veronae</i>	<i>P. fabianii</i> (A) <i>P. fabianii</i> (B) <i>P. veronae</i> (C)	A 100, B 90.1, 90.1	A 100, B 94.1, 94.1	A 100, B 100, 100	AF335971
Group 11 UWFP-363, -366, -367	Unknown <sup>e</sup>	<i>T. mucoides</i>	100	100	99.7	AF335988

<sup>a</sup> Type strains are designated by a superscript T.<sup>b</sup> API and VITEK systems, supplemented by conventional methods, were used to perform biochemical tests (2).<sup>c</sup> Identification confirmed by  $\geq 99\%$  26S rDNA gene sequence similarity with type strains and published sequences. Published sequences used for similarity comparison are given in parentheses as follows: *C. kefir* (no. U94924<sup>T</sup>), *C. krusei* (no. U76347<sup>T</sup>), *C. pararugosa* (no. U62306<sup>T</sup>), *C. albidus* (no. AF075474<sup>T</sup>), *C. diffluentis* (no. AF075502<sup>T</sup>), *C. liquefaciens* (no. AF181515<sup>T</sup>), *C. neoformans* (no. AF075484<sup>T</sup>), *P. farinosa* (no. U45739<sup>T</sup>), *P. ohmeri* (no. U45702<sup>T</sup>), *R. glutinis* (no. AF070430), *R. rubra* (no. AF189960<sup>T</sup>), *Y. lipolytica* (no. U40080<sup>T</sup>), *P. fabianii* (no. U73573<sup>T</sup>), *P. veronae* (no. U73576<sup>T</sup>), and *T. mucoides* (no. AF075515<sup>T</sup>). Sequences from this study are identical with published 26S sequences except as follows: there was 99.8% similarity to published sequences for strains of *C. pararugosa* (UWFP-348), *R. glutinis* (ATCC 32765<sup>T</sup>), and *R. rubra* (UWFP-373 and -380); 99.7% similarity for strains of *P. farinosa* (UWFP-389 and -390) and *T. mucoides* (UWFP-363, -366, and -367); and 99.5% similarity for *C. neoformans* (UWFP-360).<sup>d</sup> ITS1 and ITS2 sequence similarities were compared among clinical and type strains from our study, except for *C. kefir*, *C. krusei*, and *T. mucoides*, which were compared to published sequences: *C. kefir* (JCM1630 for ITS1 [AB011519], taxon 4911 for ITS2 [L47107], NRRL Y-8281<sup>T</sup> for 26S [U94924]), *C. krusei* (NRRL Y-5396<sup>T</sup> for 26S [U76347]), and *T. mucoides* (M9478 for ITS1 and ITS2 [AB018031], CBS7625<sup>T</sup> for 26S [AF075515]).<sup>e</sup> Clinical strains UWFP-345, -363, -366, and -367 were unidentifiable by routine automated biochemical methods. Strain UWFP-345 has 100% 26S sequence identity with *P. fabianii* and *P. veronae*. However, ITS sequence comparison clearly identified this strain as *P. fabianii*, which was concordant with its positive nitrate assimilation test. Strains UWFP-363, -366, and -367 were identified by their 100% ITS2 sequence identity with *T. mucoides* in the ITS2 study (2).

(Table 2). Nucleotide insertions or deletions in the ITS1 region were observed between type and clinical strains of *C. albicans* and *E. fibuliger*, as well as among clinical strains of *E. fibuliger* and *S. cerevisiae*. For these species, this variation accounts for the SD of  $\geq 0.5$  bp observed for the PCR product lengths measured by capillary electrophoresis (Table 1). Similar nucleotide insertions and deletions were found previously in the

ITS2 regions of type and clinical strains of *C. glabrata* and *C. tropicalis*, as well as among clinical strains of *C. tropicalis* and *S. cerevisiae* (2). Conspecific strains examined among a total of 79 isolates in this study, including 30 type strains and 5 reference strains, generally displayed  $\geq 99\%$  sequence similarity (Table 2). Together with the 53 new species-specific diagnostic alleles identified (Table 2), these data provide two important mea-

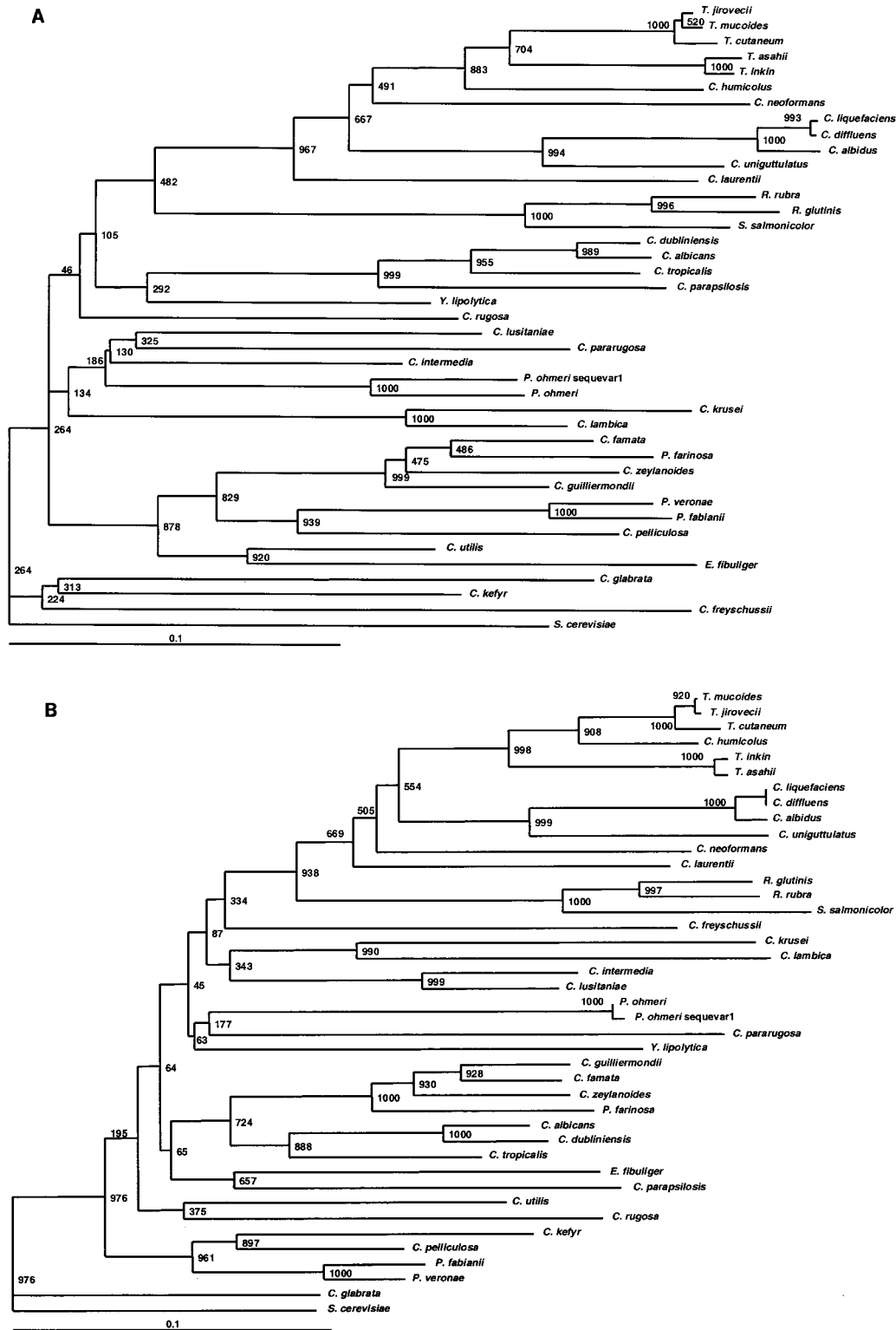


FIG. 1. (Continued on following page.)

tures of the diagnostic utility of ITS1 sequences for identifying medically important yeasts.

**Complementary ITS1 and ITS2 polymorphisms.** Twenty-one species in Table 1 comprise five groups with nearly identical ITS1 PCR product lengths, yet 11 of these species can be

identified if their ITS2 PCR product lengths are also determined: *Candida rugosa*, *Candida lusitaniae*; *Pichia ohmeri*, *Candida lambica*, *Cryptococcus laurentii*, *Cryptococcus neoformans*, *Candida utilis*, *C. dubliniensis*, *C. albicans*, *C. tropicalis*, and *Rhodotorula glutinis*. Notably, *C. rugosa*, *P. ohmeri*, *C.*

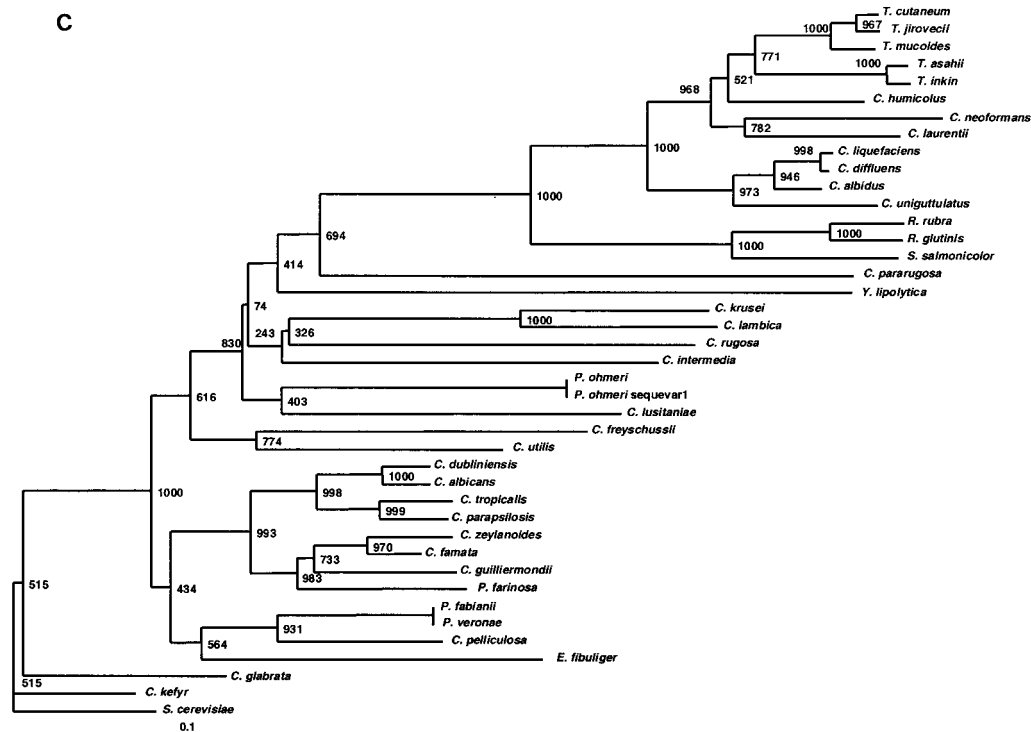


FIG. 1. (A) ITS1 sequence-based phylogenetic tree of 41 clinically significant yeast species. A consensus neighbor-joining dendrogram with 1,000 bootstrap replicates was based on 419 aligned positions of complete ITS1 sequences with adjacent partial sequences of 18S and 5.8S rRNA genes. GenBank accession numbers of sequences generated in this study are presented in Table 2. Numbers at the nodes indicate the bootstrap values. Lower bars indicate relative genetic distance. (B) ITS2 sequence-based phylogenetic tree of 42 clinically significant yeast species. Consensus neighbor-joining dendrogram with 1,000 bootstrap replicates was based on 432 aligned positions of complete ITS2 sequences with adjacent partial sequences of 5.8S and 26S rRNA genes. The sequences of *C. albicans* (no. L28817) and *C. parapsilosis* (no. U10988), as well as our sequences used for tree building, were previously published (2). Accession numbers of additional ITS2 sequences generated in this study are shown in Table 2 (footnote d). (C) 26S sequence-based phylogenetic tree of 41 clinically significant yeast species. Consensus neighbor-joining dendrogram with 1,000 bootstrap replicates was based on 565 aligned positions of the D1/D2 region of the 26S rRNA genes. Sequences of the following organisms (accession numbers in parentheses) were retrieved from GenBank: *C. albicans* (no. U45776), *C. dubliniensis* (no. U57685), *C. famata* (no. U45808), *C. freyschussii* (no. AF017242), *C. glabrata* (no. U44808), *C. guilliermondii* (no. U45709), *C. intermedia* (no. U44809), *C. kefir* (no. U94924), *C. krusei* (no. U76347), *C. lambica* (no. AF020435), *C. lusitaniae* (no. U44817), *C. parapsilosis* (no. U45754), *C. pararugosa* (no. U62306), *C. pelliculosa* (no. U74592), *C. rugosa* (no. U45727), *C. tropicalis* (no. AB034689), *C. utilis* (no. U73570), *C. zeylanoides* (no. U45832), *C. diffluens* (no. AF075502), *C. humicolus* (no. AF189836), *C. laurentii* (no. AF075469), *C. liquefaciens* (no. AF181515), *C. uniguttulatus* (no. AF075468), *E. fibuliger* (no. U40088), *P. fabianii* (no. U73573), *P. veronae* (no. U73576), *S. cerevisiae* (no. U44806), *S. salmonicolor* (no. AF070439), *T. asahii* (no. AF105393), *T. cutaneum* (no. AF075483), *T. inkin* (no. AF105396), *T. jirovecii* (no. AF105398), and *T. mucoides* (no. AF075515). Sequences of the following organisms (accession numbers in parentheses) are from this study: *C. albidus* (no. AF335982), *C. neoformans* (no. AF335984), *P. farinosa* (no. AF335974), *P. ohmeri* (no. AF335976), *P. ohmeri* sequevar 1 (no. AF335975), *R. glutinis* (no. AF335985), *R. rubra* (no. AF335986), and *Y. lipolytica* (no. AF335977).

*neoformans*, *C. laurentii*, and *C. utilis* could only be identified by using a combination of ITS1 and -2 length polymorphisms (Table 1) (2). That is, neither ITS PCR product length alone permitted unambiguous identification. ITS2 length polymorphisms alone identified 16 species of yeasts (2) compared with 19 identified by using ITS1 alone (Table 1). However, the combination of ITS1 and ITS2 length polymorphisms accurately identified 30 species of yeasts (Tables 1 and 2). The isolation rate of these 30 species in our clinical laboratory, which serves a spectrum of patients ranging from primary to tertiary care and provides reference laboratory services for five Western states, indicates that >98% of clinical isolates from our laboratory could be identified by ITS length polymorphisms.

The remaining 10 species (Table 1) included four groups that could not be distinguished by ITS region PCR product lengths (Table 3). These isolates could be readily and accurately identified by DNA sequence analysis of either ITS allele

(Table 3) or restriction endonuclease analysis of the PCR product containing both ITS1 and ITS2 (Table 3). Two exceptions were *Cryptococcus liquefaciens* and *Cryptococcus diffluens*. These closely related species (7) contained identical ITS2 alleles (Table 3) (2), provided indistinguishable *Bsa*HI and *Hinc*II restriction analyses from their ITS1 and -2 PCR products, and displayed phenotypic characteristics of *Cryptococcus albidus* by routine automated biochemical methods and therefore were only distinguishable by the DNA sequences of their ITS1 alleles (Table 3).

**Identification of unusual isolates.** In a previous study (2), we observed >98% concordance between biochemical and ITS2 genotypic identifications validated with over 400 clinical isolates. However, four clinical isolates could not be readily identified by automated biochemical systems (UWFP-345, -363, -366, and -367). Six additional clinical isolates (UWFP-348, -357, -359, -370, -373, and -380) had questionable designations:



their identity predicted by biochemical characteristics was clearly not concordant with their ITS2 genotype. The identity of these 10 isolates was established in Table 4, which includes analysis of the hypervariable D1/D2 region of the 26S rRNA gene (5, 13). These data confirm the utility of ITS sequences for correctly identifying *C. rugosa* and *C. pararugosa* (Table 4, group 3), *C. albidus*, *C. diffluens*, and *C. liquefaciens* (group 4), *R. glutinis* and *R. rubra* (group 8), *P. fabianii* and *P. veronae* (group 10), and *T. mucoides* (group 11). Notably, *P. veronae* and *P. fabianii* have identical D1/D2 hypervariable sequences and could only be identified by using ITS alleles (Table 4, group 10).

Table 4 also contains analyses of isolates with ITS1 (Table 2) or ITS2 (2) alleles that showed  $\leq 99\%$  sequence similarity between type and clinical strains or among clinical strains. Interestingly, the newly designated *C. diffluens* and *C. liquefaciens* (7) strains share 99.6, 100, and 99.5% sequence similarity at ITS1, ITS2, and the 26S rRNA hypervariable loci, respectively (Table 4, group 4). Although  $>99\%$  similarity is generally typical of conspecific strains (2, 5, 13, 27), Table 4 provides several examples where this may not be the case: *Candida kefyr* (row 1), *C. krusei* (row 2), *P. farinosa* (group 6), *P. ohmeri* (group 7), and *Y. lipolytica* (group 9). Because species delineation now encompasses the integration of both physiological and genotypic characteristics, we use the term “sequevar” to distinguish clinical isolates with the following characteristics: one diagnostic allele  $<99\%$  similar to the type strain while biochemical and two (or more) other diagnostic loci (defined as  $\geq 99\%$  similar) concur for the designation of a particular species—e.g., *P. farinosa* ITS1 sequevar 1, *P. ohmeri* ITS1 sequevar 1, and *Y. lipolytica* ITS1 sequevar 1 (Table 4, groups 6, 7, and 9). It is possible that further investigation(s) will confirm these as new species, considering the similarity of the three diagnostic loci examined for *C. diffluens* and *C. liquefaciens* (Table 4). Furthermore, ITS2 sequevars, such as for *C. kefyr* (Table 4, row 1), may also exist. Together, these data indicate that ITS analyses are useful for distinguishing closely related clinical isolates.

**Phylogenetic analysis.** In addition to discriminating among closely related species, we sought to determine if ITS1 alleles could accurately indicate the relationships among distantly related taxa. Phylogenetic trees were constructed with ITS1, ITS2, and 26S sequences (Fig. 1A, B, and C, respectively). Dendrogram topologies derived from the three sets of sequences were highly similar. Similar to the 26S tree, high bootstrap values were observed at the deep branches separating the heterogeneous ascomycetous yeasts and the different clades of the basidiomycetous yeasts. Within each clade, the relationships among species of monophyletic taxa with statistically well-supported peripheral branches were generally concordant between all three markers. Moreover, closely related species with identical 26S sequences could be separated by ITS sequences with strong bootstrap support (e.g., *P. veronae* and *P. fabiani*) and species with identical ITS2 sequences could be separated by examining their ITS1 alleles (e.g., *C. diffluens* and *C. liquefaciens*). We therefore conclude that ITS loci could be used for taxonomic placement and to infer phylogenetic relationships of previously undescribed species and novel pathogens isolated in the clinical laboratory.

## DISCUSSION

DNA-based methods have been used successfully to characterize pathogenic yeasts and provide evidence of novel species (1, 2, 5, 7, 10, 12, 14, 25, 27–29, 31). We have established an ITS database that provides accurate identification of at least 40 species of medically important yeasts. Nineteen of these can be rapidly identified by capillary electrophoresis to determine ITS1 PCR product lengths alone, compared with 16 identified by using ITS2 alone (Table 1). The use of both alleles identified 30 species (Table 1), including 4 species that could only be unambiguously designated by using both ITS alleles (see Results). Our diagnostic library of ITS length polymorphisms is validated with over 400 clinical strains (Tables 1 and 2) (2), and together, the ITS1 and -2 length polymorphisms identify  $>98\%$  of the isolates received by our clinical laboratory.

The direct sequence analysis of 79 isolates confirmed the accuracy of estimating PCR product lengths by capillary electrophoresis (2) and provided 53 new diagnostic ITS1 alleles from a well-characterized group of clinical isolates with concordant biochemical analysis-, ITS1-, and ITS2-based identifications (Table 2). ITS1 alleles showed species specificity and permitted 40 species of pathogenic yeasts to be unambiguously identified by restriction endonuclease or DNA sequence analysis of ITS1 region PCR products (Tables 2 and 3). In combination with sequences from ITS2 (55 alleles total from reference 2 and Table 2) and the D1/D2 hypervariable region of the 26S rRNA gene, unusual clinical isolates were also readily identified (Table 4). Together with the ITS1 alleles, the additional ITS2 and D1/D2 region sequences from this study provide a total of 67 new diagnostic alleles (Table 2, footnote d, and Table 4).

Data from three diagnostic loci, the D1/D2 hypervariable region of the 26S rRNA gene, ITS1, and ITS2, allowed ambiguous biochemical and genotypic designations to be clarified (Table 4). In this respect, we found ITS1 to be most useful, because it distinguished among closely related isolates with identical sequences at either the 26S rRNA gene or ITS2 loci. Similarly, others have found ITS sequences particularly useful for distinguishing among closely related yeasts (5, 12, 14, 27, 28, 31). Furthermore, our data indicating that ATCC 11066 (originally designated *Candida stellatoidea*) is *C. albicans* (Table 2) agrees with other molecular data (20) that *C. stellatoidea* does not merit species status (15, 32). Although we observed intraspecies ITS allelic variation, as has been reported previously (2, 27), conspecific strains generally demonstrated  $\geq 99\%$  sequence similarity (Table 2). This result is in agreement with studies including other loci and large numbers of diverse isolates (2, 5, 12, 14, 27, 28). However, the possibility exists for exceptions to this benchmark (Table 4), because the converse argument implies that  $\leq 99\%$  sequence similarity connotes different species: for example, *Cryptococcus diffluens* and *C. liquefaciens* demonstrate  $>99\%$  similarity at three diagnostic loci (Table 4 [see also Results and reference 7]). To facilitate accurate reporting of genetic data and to maintain the integrity of public databases for the purposes of identifying clinically significant yeasts, we use the term “sequevar” to indicate isolates that differ at one locus ( $\leq 99\%$  sequence similarity) but share two (or more) loci ( $\geq 99\%$  similarity) concordantly with biochemical and morphological indicators for a particular spe-

cies designation. Clearly, additional investigation will be necessary to accurately determine the species relationships among taxa containing such isolates.

The data in Table 4 indicate the utility of ITS sequences for distinguishing among closely related yeasts, and we also determined the relevance of ITS sequences for establishing the relationships among distantly related taxa. The phylogenetic relationships observed in the ITS1 and ITS2 trees were highly concordant with those inferred from 26S rDNA sequences (Fig. 1). ITS2 and ITS1 were better than 26S at resolving the taxonomic position of *Cryptococcus humicolus*. In agreement with our data, this species has been previously shown to belong to the *Trichosporon* clade by analysis of 18S rDNA and ITS sequences (29), as well as 26S rDNA (10). In the ascomycetous group, *Pichia* species were polyphyletic in all three dendrograms. These results are concordant with the data of Kurtzman et al., who also showed that *Pichia* species are polyphyletic and widely distributed among the ascomycetous yeasts based on 26S rDNA phylogenies (13). Thus, our data confirm that ITS1 and ITS2 regions are not only phylogenetically informative, as shown by other investigators (5, 28, 29, 31), but are also useful in distinguishing closely as well as distantly related taxa (Fig. 1). By extrapolation, an expanded ITS sequence database could be used for taxonomic placement and to infer phylogenetic relationships of previously undescribed species and novel pathogens. Ready access to these species-specific DNA sequences, amplified with universal primers, predicts that array-based hybridization schemes will become useful diagnostic tools in the clinical microbiology laboratory.

#### REFERENCES

- Baleiras Couto, M. M., B. Eijssma, H. Hofstra, J. H. in't Veld, and J. M. B. M. van der Vossen. 1996. Evaluation of molecular typing techniques to assign genetic diversity among *Saccharomyces cerevisiae* strains. *Appl. Environ. Microbiol.* **62**:41–46.
- Chen, Y. C., J. D. Eisner, M. M. Kattar, S. L. Rassouljian-Barrett, K. LaFe, S. L. Yarfitz, A. P. Limaye, and B. T. Cookson. 2000. Identification of medically important yeasts using PCR-based detection of DNA sequence polymorphisms in the internal transcribed spacer 2 region of the rRNA genes. *J. Clin. Microbiol.* **38**:2302–2310.
- Dooley, D. P., M. L. Beckius, and B. S. Jeffrey. 1994. Misidentification of clinical yeast isolates by using the updated Vitek Yeast Biochemical Card. *J. Clin. Microbiol.* **32**:2889–2892.
- el-Zaatari, M., L. Pasarelli, M. R. McGinnis, J. Buckner, G. A. Land, and I. F. Salkin. 1990. Evaluation of the updated Vitek yeast identification data base. *J. Clin. Microbiol.* **28**:1938–1941.
- Fell, J. W., T. Boekhout, A. Fonseca, G. Scorzetti, and A. Statzell-Tallman. 2000. Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. *Int. J. Syst. Evol. Microbiol.* **50**:1351–1371.
- Fenn, J. P., H. Segal, B. Barland, D. Denton, J. Whisenant, H. Chun, K. Christofferson, L. Hamilton, and K. Carroll. 1994. Comparison of updated Vitek Yeast Biochemical Card and API 20C yeast identification systems. *J. Clin. Microbiol.* **32**:1184–1187.
- Fonseca, A., G. Scorzetti, and J. W. Fell. 2000. Diversity in the yeast *Cryptococcus albidus* and related species as revealed by ribosomal DNA sequence analysis. *Can. J. Microbiol.* **46**:7–27.
- Gilfillan, G. D., D. J. Sullivan, K. Haynes, T. Parkinson, D. C. Coleman, and N. A. Gow. 1998. *Candida dubliniensis*: phylogeny and putative virulence factors. *Microbiology* **144**:829–838.
- Gleason, T. G., A. K. May, D. Caparelli, B. M. Farr, and R. G. Sawyer. 1997. Emerging evidence of selection of fluconazole-tolerant fungi in surgical intensive care units. *Arch. Surg.* **132**:1197–1201.
- Gueho, E., L. Improvisi, R. Christen, and G. S. de Hoog. 1993. Phylogenetic relationships of *Cryptococcus neoformans* and some related basidiomycetous yeasts determined from partial large subunit rRNA sequences. *Antonie Leeuwenhoek* **63**:175–189.
- Irobi, J., A. Schoofs, and H. Goossens. 1999. Genetic identification of *Candida* species in HIV-positive patients using the polymerase chain reaction and restriction fragment length polymorphism analysis of its DNA. *Mol. Cell. Probes* **14**:401–406.
- Kurtzman, C. P. 2000. Four new yeasts in the *Pichia anomala* clade. *Int. J. Syst. Evol. Microbiol.* **50**:395–404.
- Kurtzman, C. P., and C. J. Robnett. 1998. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Antonie Leeuwenhoek* **73**:331–371.
- Kurtzman, C. P., and C. J. Robnett. 1997. Identification of clinically important ascomycetous yeasts based on nucleotide divergence in the 5' end of the large-subunit (26S) ribosomal DNA gene. *J. Clin. Microbiol.* **35**:1216–1223.
- Kwon-Chung, K. J., J. B. Hicks, and P. N. Lipke. 1990. Evidence that *Candida stellatoidea* type II is a mutant of *Candida albicans* that does not express sucrose-inhibitable  $\alpha$ -glucosidase. *Infect. Immun.* **58**:2804–2808.
- Lin, D., L. C. Wu, M. G. Rinaldi, and P. F. Lehmann. 1995. Three distinct genotypes within *Candida parapsilosis* from clinical sources. *J. Clin. Microbiol.* **33**:1815–1821.
- Lodder, J. (ed.). 1970. The yeasts: a taxonomic study, 2nd ed. North-Holland Publishing Company, Amsterdam, The Netherlands.
- Lott, T. J., R. J. Kuykendall, and E. Reiss. 1993. Nucleotide sequence analysis of the 5.8S rDNA and adjacent ITS2 region of *Candida albicans* and related species. *Yeast* **9**:1199–1206.
- Masih, E. I., I. Alie, and B. Paul. 2000. Can the grey mould disease of the grape-vine be controlled by yeast? *FEMS Microbiol. Lett.* **189**:233–237.
- McCullough, M. J., K. V. Clemons, and D. A. Stevens. 1999. Molecular and phenotypic characterization of genotypic *Candida albicans* subgroups and comparison with *Candida dubliniensis* and *Candida stellatoidea*. *J. Clin. Microbiol.* **37**:417–421.
- Metzgar, D., A. van Belkum, D. Field, R. Haubrich, and C. Wills. 1998. Random amplification of polymorphic DNA and microsatellite genotyping of pre- and posttreatment isolates of *Candida* spp. from human immunodeficiency virus-infected patients on different fluconazole regimens. *J. Clin. Microbiol.* **36**:2308–2313.
- Nagahama, T., M. Hamamoto, T. Nakase, and K. Horikoshi. 1999. *Kluyveromyces nonfermentans* sp. nov., a new yeast species isolated from the deep sea. *Int. J. Syst. Bacteriol.* **49**:1899–1905.
- Oda, Y., M. Yabuki, K. Tonomura, and M. Fukunaga. 1997. A phylogenetic analysis of *Saccharomyces* species by the sequence of 18S–28S rRNA spacer regions. *Yeast* **13**:1243–1250.
- Pfaller, M. A., S. A. Messer, A. Houston, M. S. Rangel-Frausto, T. Wiblin, H. M. Blumberg, J. E. Edwards, W. Jarvis, M. A. Martin, H. C. Neu, L. Saiman, J. E. Patterson, J. C. Dibb, C. M. Roldan, M. G. Rinaldi, and R. P. Wenzel. 1998. National epidemiology of mycoses survey: a multicenter study of strain variation and antifungal susceptibility among isolates of *Candida* species. *Diagn. Microbiol. Infect. Dis.* **31**:289–296.
- Reiss, E., K. Tanaka, G. Bruker, V. Chazalet, D. Coleman, J. P. Debeaupuis, R. Hanazawa, J. P. Latge, J. Lortholary, K. Makimura, C. J. Morrison, S. Y. Murayama, S. Naoe, S. Paris, J. Sarfati, K. Shibuya, D. Sullivan, K. Uchida, and H. Yamaguchi. 1998. Molecular diagnosis and epidemiology of fungal infections. *Med. Mycol.* **36**(Suppl. 1):249–257.
- Scorzetti, G., I. Petrescu, D. Yarrow, and J. W. Fell. 2000. *Cryptococcus adeliensis* sp. nov., a xylanase producing basidiomycetous yeast from Antarctica. *Antonie Leeuwenhoek* **77**:153–157.
- Sugita, T., A. Nishikawa, R. Ikeda, and T. Shinoda. 1999. Identification of medically relevant *Trichosporon* species based on sequences of internal transcribed spacer regions and construction of a database for *Trichosporon* identification. *J. Clin. Microbiol.* **37**:1985–1993.
- Sugita, T., M. Takashima, R. Ikeda, T. Nakase, and T. Shinoda. 2000. Intraspecific diversity of *Cryptococcus laurentii* as revealed by sequences of internal transcribed spacer regions and 28S rRNA gene and taxonomic position of *C. laurentii* clinical isolates. *J. Clin. Microbiol.* **38**:1468–1471.
- Sugita, T., M. Takashima, R. Ikeda, T. Nakase, and T. Shinoda. 2000. Phylogenetic and taxonomic heterogeneity of *Cryptococcus humicolus* by analysis of the sequences of the internal transcribed spacer regions and 18S rDNA, and the phylogenetic relationships of *C. humicolus*, *C. curvatus*, and the genus *Trichosporon*. *Microbiol. Immunol.* **44**:455–461.
- Takashima, M., and T. Nakase. 2000. Four new species of the genus *Sporobolomyces* isolated from leaves in Thailand. *Mycoscience* **41**:357–369.
- Takashima, M., and T. Nakase. 1999. Molecular phylogeny of the genus *Cryptococcus* and related species based on the sequences of 18S rDNA and internal transcribed spacer regions. *Microbiol. Cult. Coll.* **15**:35–47.
- Wickes, B. L., J. E. Golin, and K. J. Kwon-Chung. 1991. Chromosomal rearrangement in *Candida stellatoidea* results in a positive effect on phenotype. *Infect. Immun.* **59**:1762–1771.
- Williams, D. W., M. J. Wilson, M. A. O. Lewis, and A. J. C. Potts. 1996. Identification of *Candida* species in formalin fixed, paraffin wax embedded oral mucosa by sequencing of ribosomal DNA. *Mol. Pathol.* **49**:M23–M28.
- Wright, W. L., and R. P. Wenzel. 1997. Nosocomial *Candida*. Epidemiology, transmission, and prevention. *Infect. Dis. Clin. N. Am.* **11**:411–425.